RAD-seq 技术及其在昆虫学研究中的应用

魏书军[1,*]，李冰艳[1,2]，曹利军[1]，朱家颖[2]

（1. 北京市农林科学院植物保护环境保护研究所，北京 100097；
2. 西南林业大学，云南省森林灾害预警与控制重点实验室，昆明 650224）

摘要：限制性酶切位点关联 DNA 测序技术(restriction-site-associated DNA sequencing, RAD-seq)是在二代测序技术基础上发展起来的一种简化基因组技术(reduced-representation genome sequencing, RRGS)。RAD-seq 利用限性核酸内切酶使基因组片段化,经过修饰后连接含标记的接头构建文库并进行测序。因其具有操作简单、实验成本低、通量高等优点,在分子生态学、进化基因组学、保护遗传学等领域得到应用。目前发展起来的 RAD-seq 技术主要包括利用稀有酶和物理打断方式进行基因组片段化的 mbRAD、利用稀有酶和常见酶组合酶切的方式进行基因组片段化的 ddRAD、利用 IIB 型限制性内切酶得到短片段文库的 2bRAD、利用同裂酶和 Illumina 试剂盒建库的 ezRAD 和完全利用 Nextera 试剂盒建库的 nextRAD。本文对每种 RAD-seq 技术的特点及其在昆虫种群遗传学、群落生态学、物种界定、系统发育和遗传连锁图谱构建等研究领域的应用进行了综述。

关键词：RAD-seq；昆虫学；种群遗传学；物种界定；系统发育；遗传连锁图谱

中图分类号：Q966 **文献标识码**：A **文章编号**：0454-6296(2016)07-0767-08

RAD-seq and its application in entomological research

WEI Shu-Jun[1,*], LI Bing-Yan[1,2], CAO Li-Jun[1], ZHU Jia-Ying[2] (1. Institute of Plant and Environmental Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China; 2. Key Laboratory of Forest Disaster Warning and Control of Yunnan Province, Southwest Forestry University, Kunming 650224, China)

Abstract: Restriction-site-associated DNA sequencing (RAD-seq) is a technology of reduced-representation genome sequencing (RRGS), which is developed with next-generation sequencing. In terms of RAD-seq, genome is digested by restriction endonuclease. The digested fragments are ligated with barcoded adapters, and used to construct library for sequencing. With the advantages of simple operation, low experimental cost, and high throughput, RAD-seq is widely used in molecular ecology, evolutionary genomics and conservation genetics. Recently, five different RAD-seq methods including mbRAD, ddRAD, 2bRAD, ezRAD and nextRAD have been developed. Of them, mbRAD uses a low frequency cutter and physical shearing to make genome into fragments. ddRAD uses a low frequency cutter and a high frequency cutter to digest genome. 2bRAD relies on IIB-type restriction endonuclease to excise genomic sequences into short fragments. ezRAD uses isochimozer enzymes and Illumina library preparation kit. nextRAD entirely uses Nextera DNA library preparation kit. In this article, we reviewed the characteristics of each RAD-seq method and its application in entomology, such as population genetics, community ecology, species delimitation, phylogenetic reconstruction and construction of genetic linkage map.

Key words: RAD-seq; entomology; population genetics; species delimitation; phylogeny; genetic linkage map

---

基金项目：北京市自然科学基金项目(6162010)；国家自然科学基金项目(31472025)；国家重点基础研究发展计划项目(2013CB127600)
作者简介：魏书军，男，1981 年 12 月生，山东日照人，博士，副研究员，研究方向为昆虫分子生态学，E-mail：shujun268@163.com
* 通讯作者 Corresponding author, E-mail：shujun268@163.com
收稿日期 Received：2016-04-21；接受日期 Accepted：2016-06-05

二代测序(next-generation sequencing, NGS)技术的发展使得人们能够以较低的价格快速测定生物基因组和转录组,推动了分子生态学、进化基因组学、保护遗传学等领域发生突破性的变革(Arnold *et al.*, 2013; Andrews *et al.*, 2016),生命科学进入一个新的时代(Ellegren, 2014)。在二代测序技术的基础上出现了简化基因组测序技术(reduced-representation genome sequencing, RRGS)。该技术运用生物信息学方法设计特定的标记标签,对富集的特异性片段进行高通量测序,以测定的部分基因组序列代表目标物种全基因组信息。常见的 RRGS 有限制性酶切位点关联 DNA 测序(restriction-site-associated DNA sequencing, RAD-seq)、基于测序的基因分型(genotyping by sequencing, GBS)、简化代表文库(reduced-representation libraries, RRL)和简化多态序列复杂性(complexity reduction of polymorphic sequences, CRoPS)等(Andrews and Luikart, 2014)。

RAD-seq 是 RRGS 中发展最快、应用最广泛的一种(Davey *et al.*, 2011; Schweyen *et al.*, 2014)。它利用限制性核酸内切酶(restriction endonuclease, RE)对物种基因组进行酶切,回收一定大小的片段后,构建文库并进行测序,具有通量高、成本低、实验周期较短且无需参考基因组等优点(Davey *et al.*, 2013)。相比于一代测序技术(Sanger sequencing)和基因组重测序技术(whole genome re-sequencing), RAD-seq 不仅能用于系统进化模式的分析和群体基因组扫描,而且在分子标记的开发、遗传图谱的构建、重要经济性状的数量性状基因座(quantitative trait locus, QTL)定位、群体遗传结构分析、群落构建机制、连锁不平衡作图(又叫联合作图, associating mapping)等方面得到应用(Gugerli *et al.*, 2013; Andrews and Luikart, 2014; Ellegren, 2014; Hand *et al.*, 2015)。

1　RAD-seq 的基本流程

RAD-seq 的基本流程包括基因组 DNA 提取、酶切、连接接头、片段回收、文库构建、测序和数据分析等(图 1)。该过程需一定量的基因组 DNA,通常在 50 ~ 100 ng,特别要注意起始 DNA 的质量,过度的降解会导致 RAD 标记和单核苷酸多态性(single nucleotide polymorphism, SNP)位点的减少(Graham *et al.*, 2015)。进行酶切时要选择识别位点在基因组上均匀分布的限制性内切酶,以保证能够获得足够数量的 RAD 标记。对于已知基因组序列信息的物种,可先用电子酶切的方法进行模拟以选择合适的限制性内切酶,常用的程序有 DDsilico (Rašić *et al.*, 2014)和 Digital_RADs.py(https://github.com/BU-RAD-seq/Digital_RADs)等。酶切的主要目的是使基因组片段化,然后在两端连接含标记的接头,通过凝胶电泳回收和磁珠吸附等方法进行片段大小选择,完成文库的构建。建库完成后即可上机测序。目前常用的 RAD-seq 测序平台有 Illumina GAIIx, Illumina HiSeq 2000 和 Illumina HiSeq 2500 等。

测序完成后对测序结果进行分析,可以通过 Stacks (http://creskolab.uoregon.edu/stacks/) (Catchen *et al.*, 2013)软件包实现标记和酶切位点的识别和修正、read 的聚类等过程。该软件包可以与 MySQL 数据库关联,将每一步的分析结果输入数据库,从而可以在网络浏览器中查看,也方便对数据进行前期处理。

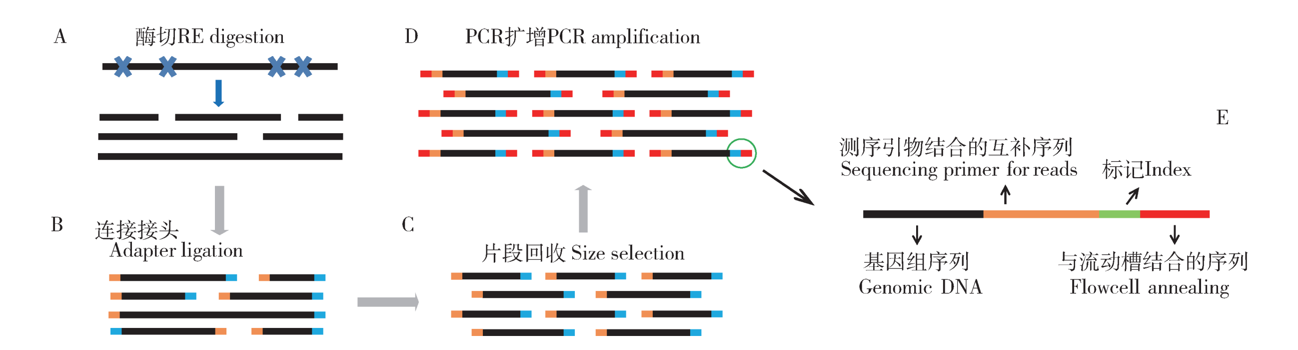


图 1　RAD-seq 技术的基本流程

Fig. 1　General steps of RAD-seq method

A: 限制性内切酶对基因组 DNA 进行酶切 Restriction endonuclease (RE) digestion of genomic DNA; B: 在酶切产物上连接接头 Adapter ligation to the digested fragments; C: 对特定长度片段进行回收 Size selection; D: PCR 扩增回收片段 PCR amplification of the selected segments; E: 文库一端的结构示意图,即对 D 中绿色圆圈区域的放大 Structure of one side of the library, corresponding to the green circle area in figure D.

2　各种 RAD-seq 技术的特点

目前已发表的 RAD-seq 技术主要有 5 种，包括 mbRAD，ddRAD，2bRAD，ezRAD 和 nextRAD（Puritz *et al*.，2014；Andrews *et al*.，2016），每种方法各有特点（表 1），研究者可根据实验目的和实验条件进行选择。

2.1　mbRAD（original RAD-seq）

mbRAD 是最原始的一种 RAD-seq 技术。该方法的步骤较为复杂，仅用一种在基因组中识别位点较少的限制性内切酶进行酶切，在酶切片段两端连接含标记的接头 P1 后进行物理打断（如超声波），打断后再连接接头 P2 构建文库，最后进行 PCR 扩增（图 2：A）（Miller *et al*.，2007；Baird *et al*.，2008）。

虽然 mbRAD 是 5 种方法中技术和步骤最为复杂的一种，但是由于物理打断的随机性，插入较长片段的可能性较大，一方面可以鉴定 PCR 扩增引入的重复拷贝，另一方面使得在从头组装时可获得比其他方法更长的序列，因此在基因功能鉴定和基因本体（gene ontology）分析中得到应用。同时，这种随机性也造成了不同长度的片段的测序深度不同，造成测序结果不均衡（Davey *et al*.，2013）。

2.2　ddRAD（double-digest RAD）

ddRAD 通过一种识别位点较少的稀有酶和一种识别位点较多的常见酶对基因组 DNA 进行双酶切，在片段两端分别连接 P1 接头（含一组 barcode）和 P2 接头（为 Y-型结构），将样品合并后进行片段大小选择，然后通过 PCR 引物在第 2 个酶切位点附近加入一组 index（区别于 barcode），在这一过程中 Y-型结构还确保了两端同为 P2 接头的片段无法扩增进入到文库（图 2：B）（Peterson *et al*.，2012）。

与 mbRAD 相比，该方法没有物理打断的过程。由于物理打断的随机性比较大，相比之下用两种限制性内切酶酶切的目的性较强，使得其对 DNA 文库的筛选更为严格。其次，研究者可以通过对不同的限制性内切酶的组合以及对片段大小选择范围的优化，使单个个体在有限的测序通量中得到合适的测序深度，从而利用较少的测序量就可得到可靠的结论。ddRAD 采用双重标记减少了接头和引物的合成数量，从而降低了试验成本。如，对 96 个样品进行区分时，mbRAD 需要 96 种标记，而 ddRAD 只需 8 种 barcode 和 12 种 index，即 8 对接头和 12 条引物。但是 ddRAD 需要更高质量的基因组 DNA。片段大小选择的不准确性容易导致个体间测序位点的差异较大，虽然可以使用准确的片段大小选择方法，如使用全自动核酸电泳和片段回收系统 Pippin Prep（Sage Science，Inc.，USA），但同时也会降低基因的多样性。当限制酶的识别位点缺失或具多态性时，将导致酶切错误，容易产生等位基因丢失［又叫等位基因脱扣（allele dropout，ADO）］，降低对遗传多样性的估计。由于 ddRAD 方法需用两种限制酶，所以结果更易受 ADO 的影响（Arnold *et al*.，2013）。因此，在对敏感种群进行基因分析时需考虑该因素（Puritz *et al*.，2014）。另外，PCR 扩增的过程容易产生 PCR 重复拷贝，导致测序不均衡，扩大了纯合子基因位点所占的比例。目前可以通过在粘性末端和接头之间插入一个短的简并位点区（degenerate base region，DBR）（5′- NNNNNIIICC-3′）序列来检测 PCR 重复拷贝（Schweyen *et al*.，2014），这种方法也可以应用于单末端 RAD-seq（Baird *et al*.，2008）和 2bRAD（Wang *et al*.，2012）。

2.3　2bRAD

2bRAD 建库的特别之处在于选用 IIB 型限制性内切酶，如 *Bsa*Ⅺ，*Alf*Ⅰ和 *Bae*Ⅰ，这类酶分别对靠近识别位点的上游和下游序列进行酶切，获得的片段长度基本一致，约 33 ~ 36 bp。酶切后进行接头的连接，需用不同标记来区分不同个体，最后 PCR 扩增并进行电泳，切胶回收目的条带完成建库（图 2：C）（Wang *et al*.，2012）。

相比于其他几种 RAD 方法，2bRAD 过程比较简单，回收的片段很小，选用 50 bp 单末端的 Illumina 测序试剂盒测序即可。这种方法降低了基因组的复杂度，可快速有效地得到成千上万 SNPs 标记。尤其是在基因分型和 QTL 作图中，测定较少的序列就可以得到理想的结果。这种短的片段虽然足够代表基因组的特异性，但在基因组重复度高的物种中效果较差。为此，2bRAD 片段不能用于构建基因组重叠群，也不能用于诸如不同物种之间的遗传距离较大的图谱构建。

2.4　ezRAD

ezRAD 是用识别位点高频且相同的一种或多种限制性内切酶，即同裂酶，使基因组片段化，产生的片段可用 Illumina TruSeq 试剂盒建库，只需再合并文库进行片段大小选择。最后的 PCR 过程是否需要可根据建库试剂盒而定（图 2：D）（Toonen *et al*.，2013）。

该方法最大的优点就是不需要研究者自己设计

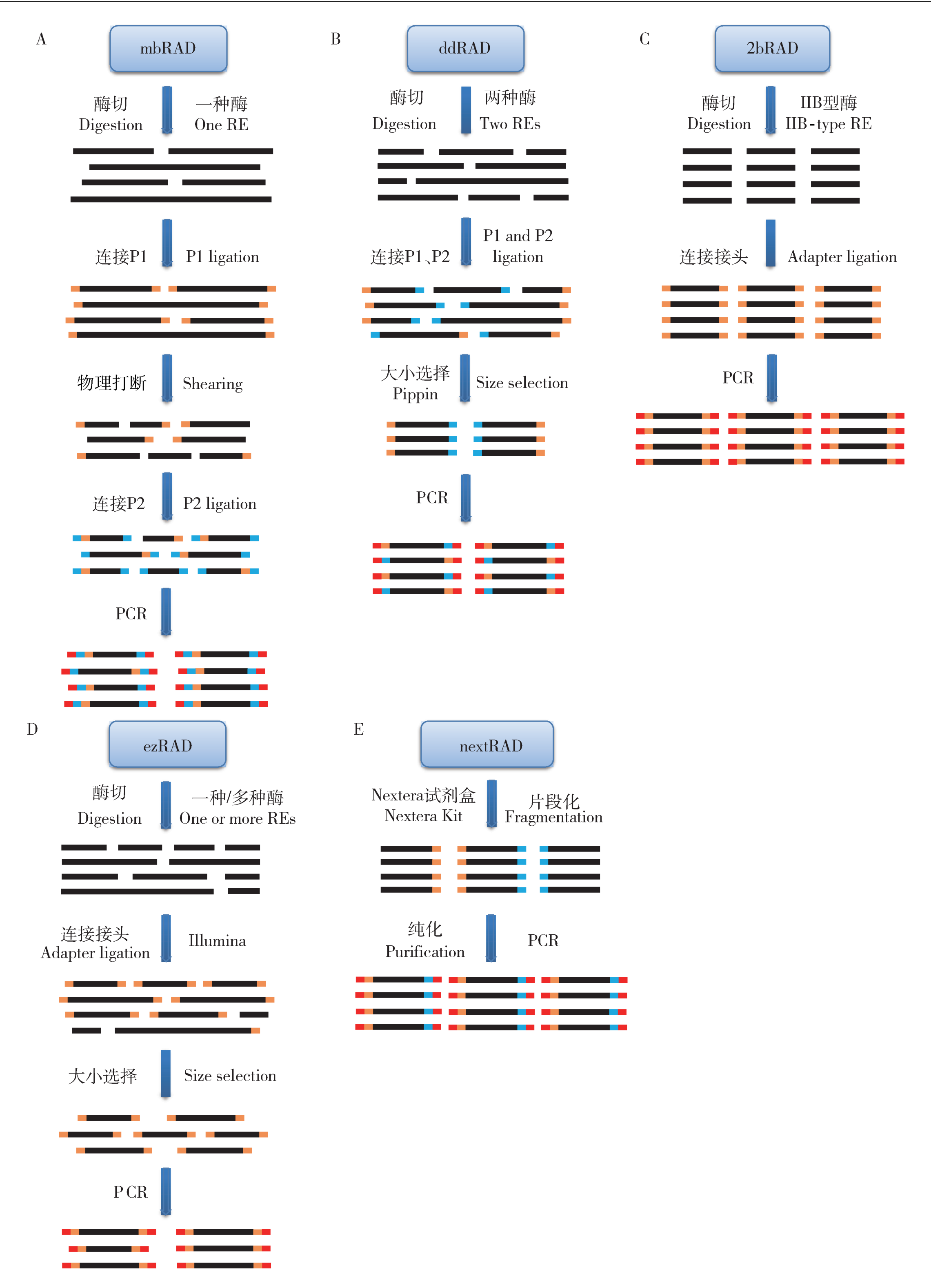


图 2 5 种 RAD-seq 技术的原理示意图

Fig. 2 Principle of 5 RAD-seq methods

A：mbRAD，使用稀有限制性内切酶和物理打断组合方法对基因组 DNA 进行片段化 mbRAD，genomic DNA fragmentation by a low frequency cutter and physical shearing；B：ddRAD，使用稀有限制性内切酶和常见限制性内切酶组合酶切基因组 DNA ddRAD，genomic DAN fragmentation by a low frequency cutter and a high frequency cutter；C：2bRAD，使用 IIB 型限制性内切酶酶切基因组 DNA 2bRAD，genomic DNA fragmentation by IIB-type restriction endonuclease；D：ezRAD，使用同裂酶酶切基因组 DNA，再使用 Illumina 试剂盒建库 ezRAD，genomic fragmentation by isochimozer enzymes and library construction by Illumina TruSeq library Prep Kit；E：nextRAD，使用 Nextera 试剂盒建库 nextRAD，library construction by Nextera DNA Library Prep Kit. RE：限制性内切酶 Restriction endonuclease.

接头及引物序列，直接使用 Illumina TruSeq 试剂盒建库，同时也节省了合成接头和引物所需的费用。如果选用 PCR-Free TruSeq 试剂盒，还可省去最后的 PCR 过程，从而避免了该过程产生的实验误差。虽然 Illumina TruSeq 建库试剂盒的成本较高，但在器具（如 96 孔磁力架）和试剂（如磁珠、adapter 和引物等）方面又节省了费用。因此，ezRAD 技术适合仪器设备不齐、缺乏操作经验、建库数量不多的实验室使用。测序是从测序引物结合区下游开始进行的，ezRAD 文库的测序引物结合区下游为目的序列，ddRAD 文库在测序引物结合区下游为 barcode 和目的序列，而 ezRAD 从酶切位点开始时，相邻不同簇间 read1 上的前几个碱基均相同，而 Illumina 测序仪是通过 read1 上最开始的 4 ~ 5 个核苷酸来区别相邻簇，就会错误地把 2 个不同的 DNA 片段产生的簇归为一个簇，给出一个测序结果，降低测序质量。

2.5　nextRAD

nextRAD 建库的整个过程都是利用 Nextera DNA Sample Pre Kits（Illumina, Inc.）完成。首先，提取的基因组 DNA 经过一个 Tagmentation 过程，即在有接头序列的转座子的作用下使 DNA 片段化并连接好接头，纯化后进行 PCR 扩增，补齐片段两端的接头，并引入 index 和 flowcell 序列（图 2：E）（Russello *et al*., 2015）。和 ezRAD 一样，利用试剂盒建库操作简单，也不需要各种仪器。Nextera 试剂盒要求的 DNA 起始量比较低，50 ng 即可，所以 nextRAD 方法比较适合对小型基因组、扩增子等进行研究，特别是一些只能采用非侵入性方式收集样本（如搜集毛发）的濒危物种的研究。另外，使用这种方法还需特别注意 DNA 降解对实验结果的影响，因为降解的片段在 Tagmentation 过程中无法成功连接接头。

3　RAD-seq 在昆虫学研究中的应用

昆虫是世界上种类最为繁多的类群，除了黑腹果蝇 *Drosophila melanogaster*（Hoskins *et al*., 2007）、西方蜜蜂 *Apis mellifera*（Crozier and Crozier, 1993; Weinstock, 2006; Haddad *et al*., 2015）、家蚕 *Bombyx mori*（Mita *et al*., 2004）和小菜蛾 *Plutella xylostella*（Jouraku *et al*., 2013）等有限物种的全基因组被测定外，更多物种的基因组序列尚未被测定。RAD-seq 技术的最大优点就是可以对没有参考基因组的非模式物种进行基因组层面的研究。因此，该技术将会在昆虫学研究中发挥重要作用。目前已有报道利用 RAD-seq 进行昆虫生物地理学和种群遗传学、物种界定、系统发育重建、遗传连锁图谱构建等方面的研究。

3.1　生物地理学与种群遗传学

系统发生生物地理学（phylogeography，又称“亲缘地理学”）主要研究控制亲缘谱系地理分布的原理和过程，通过分子标记揭示物种现有种群的遗传结构，并运用系统发生学思想研究亲缘关系密切的种间或种内基因谱系现有分布格局的形成过程和机制（Avise *et al*., 1987）。传统分子标记如部分线粒体基因、核基因序列和同工酶等由于受到检测方法的限制因而数据量相对较少，而 RAD-seq 可以获得大量的分子标记（SNP），为研究生物在冰期的迁移动态（Emerson *et al*., 2010）、推测近源物种间的分化时间（Papadopoulou and Knowles, 2015）等提供新的途径。

Emerson 等（2010）用 *Sbf* Ⅰ 限制性内切酶对北美瓶草蚊 *Wyeomyia smithii* 的 21 个种群的基因组进行 RAD-seq 分析，共得到 2 750 万条 reads 和 3 741 个 SNPs，并利用这些 SNP 位点构建系统发育关系来揭示种群遗传结构和进化演变方向，结果将这 21 个种群分为南部种群和北部种群，其中南部种群分为墨西哥湾（Gulf Coast）和北卡罗来纳海岸（North Carolina Coast）两个分支，北部种群分为阿帕拉契亚山脉（Appalachian Mountain）和北部（Northern）两个分支。该研究还推测阿帕拉契亚山脉分支是由于距今约 20 000 年的劳伦泰德冰盖的退缩、气候的改变逐渐形成的，并根据北美瓶草蚊寄主的分布以及盛行风的特点推测北部分支是由阿帕拉契亚分支迁飞演化形成的。

为了揭示遗传分化时间与地理分布的关系，Papadopoulou 和 Knowles（2015）对来自维尔京群岛 10 个不同岛屿的一种不会飞行的加勒比蟋蟀 *Amphiacusta sanctaecrucis* 种群进行研究，并通过对 136 个个体的 32 000 个 SNPs 研究分析，发现它们的遗传分化模式同岛屿的周期性连通和分离有关，并推测出其遗传分化时间为 75 ~ 130 ka（thousands year ago）。

3.2　物种界定

昆虫分类的技术和手段很多，传统方法包括根据形态学、生态学、生理学等特征进行鉴别、比较和分类。随着技术的不断发展，辅以基因和染色体等标记的分析，能够更准确地对昆虫进行种类界定。

Fischer 等(2015)结合 RAD-seq 和形态学方法，重新界定了 2 种来自卡卡梅加雨林的木盲切叶蚁 *Carebara* Westwood。过去，仅根据形态学特征将能够阻塞洞口的大头兵蚁归为 *C. elmenteitae*，将无这种特点的兵蚁和工蚁归为 *C. thoracica*。经过 RAD-seq 测序结果进行最大似然分析构建进化树，明确了 *C. elmenteitae* 和 *C. thoracica* 为同一种，并将其归为 *C. phragmotica* sp. n.。

表 1　5 种 RAD-seq 方法的特征比较

Table 1　Comparison of five RAD-seq methods

特征 Trait	mbRAD	ddRAD	2bRAD	ezRAD	nextRAD
限制性内切酶的数量 Number of restriction endonucleases	1	2	1	≥1	2
限制性内切酶的种类 Type of restriction endonucleases	稀有酶 A low-frequency cutter	稀有酶和常见酶 A low-frequency cutter and a high-frequency cutter	IIB 型酶 IIB-type RE	同裂酶 Isoschizomer	Nextera 试剂盒 Nextera DNA Library Prep Kit
物理打断 Physical shearing	有 Yes	无 No	无 No	无 No	无 No
DNA 起始量 Starting amount of genomic DNA	50 – 100 ng	50 – 100 ng	50 – 100 ng	50 – 100 ng*	< 50 ng
插入片段大小范围 Target insert size	200 – 600 bp	自定义 Customization	33 – 36 bp	自定义 Customization	250 – 300 bp
片段大小选择 Size selection	有 Yes	有 Yes	无 No	有 Yes	无 No
步骤数 Number of steps	6	4	4	4 – 6	2
Contigs > 200 bp	是 Yes	否 No	否 No	有些片段 Some	是 Yes
从头组装片段长度 Length of *de novo* contigs	长 Long	中等 Medium	短 Short	中等 Medium	中等 Medium
PCR 重复拷贝的检测 Detection of PCR duplicates	可 Yes	不可 No	不可 No	不可 No	不可 No
建库成本 Library cost	高 High	很低 Very low	低 Low	中等 Medium	中等 Medium

* ezRAD 选用 PCR-free 试剂盒时需 1 ~ 2 μg 的 DNA 起始量。For PCR-free Kit, the starting genomic DNA for ezRAD is 1 – 2 μg.

3.3　系统发育

Rougerie 等(2015)结合 DNA 条码技术和 RAD-seq 调查彗尾蛾(comet moths)和水青蛾(moon moths)的多样性，包含 3 个属(*Actias*, *Argema* 和 *Graellsia*)共 48 个样本，用 *Pst*Ⅰ进行酶切最终测序得到 3 550 个 reads 和 3 921 个 SNP 位点，并在此基础上构建系统发育树，能够很好地反映种间关系，并解决了建树时 *Actias sinensis* 和 *A. felicis* 两个种团间的冲突。

Ebel 等(2015)利用 RAD-seq 重建了新热带区的悌蛱蝶属 *Adelpha* 与其温带姐妹属线蛱蝶属 *Limenitis* 的系统发育关系，结果表明在幼虫及成虫期的自然选择导致了悌蛱蝶属 *Adelpha* 的快速适应的多样化，也导致了其寄主的多样性。

3.4　遗传连锁图谱构建

Baxter 等(2011)对小菜蛾 *Plutella xylostella* 的回交系进行 RAD-seq 分析，将 3 177 条母系遗传等位基因定位到 31 条染色体上，成功构建了小菜蛾的遗传连锁图谱，并且与家蚕 *B. mori* 的染色体进行了比较基因组分析。

3.5　其他方面

Rašić等(2014)采用 ddRAD 技术研究分别来自印度尼西亚、澳大利亚、越南和巴西等 4 个地理种群的埃及伊蚊 *Aedes aegypti*，利用 DDsilico 和 DDemux 两个程序分别模拟酶切和分析测序结果，得到最佳酶切组合 *Nla*Ⅲ 和 *Mlu*CI。在 18 147 个 SNPs 中，47.9% 位于基因间隔区，17.7% 位于内含子，17.8% 位于外显子，并基于埃及伊蚊基因组研究结果将 66.4% 的 SNPs 定位在 3 条染色体上，而且证明了通过 ddRAD 获得 SNP 的结果比微卫星能更清楚地揭示不同地理种群之间的隔离。

Grenier 等(2015)也通过 ddRAD 技术分析了 *D. melanogaster* 的 5 个地理种群的差异，包括中国、希腊、荷兰、塔斯马尼亚岛和津巴布韦。一方面，从得到的 5.78 M SNPs 和 971 K 小片段插入缺失(small indel)分析，结果表明，每条染色体上约有 50% 的突变在多个种群中出现，而且非洲种群含有大部分的突变；另一方面，从种群的遗传距离和种群结构分析说明非洲种群与其他种群产生明显的分离。因此，由于地理结构特点和高水平的遗传多样性，在研究其他种群的基因组特征时可以将非洲种群的基因组作为参考。

4　小结与展望

目前,RAD-seq 技术不仅在开发分子标记、构建遗传图谱、分析种群遗传结构等方面得到应用,还可用于分析转录组和基因结构等。这种方法是在二代测序技术的基础上发展起来的,因其实验操作简便、周期较短、成本低、无需参考基因等优点而得到广泛应用并不断成熟。由于 RAD-seq 具有的特殊优点,它为我们提供了一种研究昆虫遗传差异的有效方法。可基于该技术构建遗传图谱来对亲缘关系较近的物种进行对比,获得更多的遗传信息,可以更准确地解决近缘物种之间的进化关系。其次,RAD-seq 为非模式昆虫物种的基因结构和遗传多样性研究提供了一种简便方法。特别是可以用该技术对不同地理种群进行研究,揭示各地理种群的进化历史。同时,还可以阐释各地理种群分布与地理条件之间的关系以及各地理种群之间的亲缘关系、遗传分化时间、冰期或间冰期种群动态特征等。随着技术的不断优化,RAD-seq 在基因组结构和种群遗传分析等方面更具广阔的应用前景。

参考文献(References)

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA, 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.*, 17(2): 81–92.

Andrews KR, Luikart G, 2014. Recent novel approaches for population genomics data analysis. *Mol. Ecol.*, 23(7): 1661–1667.

Arnold B, Corbett-Detig RB, Hartl D, Bomblies K, 2013. RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Mol. Ecol.*, 22(11): 3179–3190.

Avise JC, Arnold J, Ball RM, Bermingham E, Lamb T, Neigel JE, Reeb CA, Saunders NC, 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.*, 18: 489–522.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA, 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, 3(10): e3376.

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD, Blaxter ML, 2011. Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE*, 6(4): e19315.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA, 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.*, 22(11): 3124–3140.

Crozier RH, Crozier YC, 1993. The mitochondrial genome of the honeybee *Apis mellifera*: complete sequence and genome organization. *Genetics*, 133(1): 97–117.

Davey JW, Cezard T, Fuentes-Utrilla P, Eland C, Gharbi K, Blaxter ML, 2013. Special features of RAD Sequencing data: implications for genotyping. *Mol. Ecol.*, 22(11): 3151–3164.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML, 2011. Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.*, 12(7): 499–510.

Ebel ER, DaCosta JM, Sorenson MD, Hill RI, Briscoe AD, Willmott KR, Mullen SP, 2015. Rapid diversification associated with ecological specialization in Neotropical *Adelpha* butterflies. *Mol. Ecol.*, 24(10): 2392–2405.

Ellegren H, 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol. Evol.*, 29(1): 51–63.

Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, Holzapfel CM, 2010. Resolving postglacial phylogeography using high-throughput sequencing. *Proc. Natl. Acad. Sci. USA*, 107(37): 16196–16200.

Fischer G, Azorsa F, Garcia FH, Mikheyev AS, Economo EP, 2015. Two new phragmotic ant species from Africa: morphology and next-generation sequencing solve a caste association problem in the genus *Carebara* Westwood. *ZooKeys*, (525): 77–105.

Graham CF, Glenn TC, McArthur AG, Boreham DR, Kieran T, Lance S, Manzon RG, Martino JA, Pierson T, Rogers SM, Wilson JY, Somers CM, 2015. Impacts of degraded DNA on restriction enzyme associated DNA sequencing (RADSeq). *Mol. Ecol. Resour.*, 15(6): 1304–1315.

Grenier JK, Arguello JR, Moreira MC, Gottipati S, Mohammed J, Hackett SR, Boughton R, Greenberg AJ, Clark AG, 2015. Global diversity lines – a five-continent reference panel of sequenced *Drosophila melanogaster* strains. *G3 (Bethesda)*, 5(4): 593–603.

Gugerli F, Brandl R, Castagneyrol B, Franc A, Jactel H, Koelewijn HP, Martin F, Peter M, Pritsch K, Schroder H, Smulders MJM, Kremer A, Ziegenhagen B, Evoltree JC, 2013. Community genetics in the time of next-generation molecular technologies. *Mol. Ecol.*, 22(12): 3198–3207.

Haddad NJ, Loucif-Ayad W, Adjlane N, Saini D, Manchiganti R, Krishnamurthy V, AlShagoor B, Batainh AM, Mugasimangalam R, 2015. Draft genome sequence of the Algerian bee *Apis mellifera intermissa*. *Genom. Data*, 4: 24–25.

Hand BK, Lowe WH, Kovach RP, Muhlfeld CC, Luikart G, 2015. Landscape community genomics: understanding eco-evolutionary processes in complex environments. *Trends Ecol. Evol.*, 30(3): 161–168.

Hoskins RA, Carlson JW, Kennedy C, Acevedo D, Evans-Holm M, Frise E, Wan KH, Park S, Mendez-Lago M, Rossi F, Villasante A, Dimitri P, Karpen GH, Celniker SE, 2007. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science*, 316(5831): 1625–1628.

Jouraku A, Yamamoto K, Kuwazaki S, Urio M, Suetsugu Y, Narukawa

J, Miyamoto K, Kurita K, Kanamori H, Katayose Y, Matsumoto T, Noda H, 2013. KONAGAbase: a genomic and transcriptomic database for the diamondback moth, *Plutella xylostella*. *BMC Genomics*, 14: 464.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA, 2007. Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res.*, 17(2): 240 – 248.

Mita K, Kasahara M, Sasaki S, Nagayasu Y, Yamada T, Kanamori H, Namiki N, Kitagawa M, Yamashita H, Yasukochi Y, Kadono-Okuda K, Yamamoto K, Ajimura M, Ravikumar G, Shimomura M, Nagamura Y, Shin IT, Abe H, Shimada T, Morishita S, Sasaki T, 2004. The genome sequence of silkworm, *Bombyx mori*. *DNA Res.*, 11(1): 27 – 35.

Papadopoulou A, Knowles LL, 2015. Genomic tests of the species-pump hypothesis: recent island connectivity cycles drive population divergence but not speciation in Caribbean crickets across the Virgin Islands. *Evolution*, 69(6): 1501 – 1517.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE, 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5): e37135.

Puritz JB, Matz MV, Toonen RJ, Weber JN, Bolnick DI, Bird CE, 2014. Demystifying the RAD fad. *Mol. Ecol.*, 23(24): 5937 – 5942.

Rašić G, Filipović I, Weeks AR, Hoffmann AA, 2014. Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*. *BMC Genomics*, 15: 275.

Rougerie R, Cruaud A, Rasplus JY, 2015. A tale of long tails: combining DNA barcoding and RAD-sequencing to investigate the diversification of comet- and moon-moths (Lepidoptera, Saturniidae). *Genome*, 58: 272 – 273.

Russello MA, Waterhouse MD, Etter PD, Johnson EA, 2015. From promise to practice: pairing non-invasive sampling with genomics in conservation. *PeerJ*, 3: e1106.

Schweyen H, Rozenberg A, Leese F, 2014. Detection and removal of PCR duplicates in population genomic ddRAD studies by addition of a degenerate base region (DBR) in sequencing adapters. *Biol. Bull.*, 227(2): 146 – 160.

Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR, Bird CE, 2013. ezRAD: a simplified method for genomic genotyping in non-model organisms. *PeerJ*, 1: e203.

Wang S, Meyer E, McKay JK, Matz MV, 2012. 2b-RAD: a simple and flexible method for genome-wide genotyping. *Nat. Methods*, 9(8): 808 – 810.

Weinstock GM, Robinson GE, Gibbs RA, Worley KC, Evans JD, Maleszka R, Robertson HM, Weaver DB, Beye M, Bork P, 2006, Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature*, 443(7114): 931 – 949.

（责任编辑：袁德成）